da\_410\_project2\_grahn

Jason Grahn

1/14/2019

## Download airpoll.txt. In this problem, we will only focus on the first 16 observations (cities). Read the data into R (as a data frame) and name the data as airpol.full. Then use the following code to “extract” the first 16 observations.

## Use R to perform the following analysis on the subset data airpol.data.sub. Make sure you include clear headings, command lines, and relevant output/results.

### a) Calculate the sample covariance matrix and the sample correlation matrix.

### b) Calculate the distance matrix for these observations (after scaling the variables by dividing each variable by its standard deviation). Describe some of the most similar pairs of cities and some of the most different pairs of cities, giving evidence from the distance matrix.

### c) Display a plot that will help assess whether this data set comes from a multivariate normal distribution. What is your conclusion based on the plot?